



KOREAN PATENT ABSTRACTS

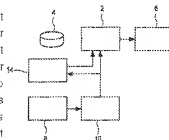
(11)Publication number: 1020020059706 A
 (43)Date of publication of application: 13.07.2002

(21)Application number: 1020027005913 (71)Applicant: KONINKLIJKE PHILIPS
 (22)Date of filing: 07.05.2002 ELECTRONICS N.V.
 (30)Priority: 08.09.2000 EP2000 00203098 (72)Inventor: BARBIERI MAURO
 (51)Int. Cl. G11B 27/10

(54) AN APPARATUS FOR REPRODUCING AN INFORMATION SIGNAL STORED ON A STORAGE MEDIUM

(57) Abstract

An apparatus for reproducing an information signal stored on a first storage medium (4). The apparatus comprises a reading unit (2) for reading the information signal from the first storage medium, an output unit (6) for supplying the information signal to a display unit, an user controllable input unit (8) for receiving commands to enable an user to access the information signal. The user controllable input unit is adapted to receive a first command at an instant. The apparatus further comprises a unit (10) for controlling the reading unit to start reading the information signal from said storage medium at a second position in the information signal in response to said first command, the information signal at said second position having features showing a similarity with features of the information signal at a first position read at said instant of receiving said first command, or a features of an information signal read prior to said instant.



copyright KIPO & WIPO 2007

Legal Status

Date of request for an examination (20050905)
 Notification date of refusal decision (00000000)
 Final disposal of an application (rejection)
 Date of final disposal of an application (20080925)
 Patent registration number ()
 Date of registration (00000000)
 Number of opposition against the grant of a patent ()
 Date of opposition against the grant of a patent (00000000)
 Number of trial against decision to refuse (2008101005973)
 Date of requesting trial against decision to refuse (20080717)
 Date of extinction of right ()

공개특허 제2002-59706호(2002.07.13.) 1부.

특2002-0059706

(19) 대한민국특허청(KR)

(12) 공개특허공보(A)

(51) Int. Cl.⁷

G11B 27/10

(11) 공개번호 특2002-0059706

(43) 공개일자 2002년07월13일

(21) 출원번호 10-2002-7065813

(22) 출원일자 2002년06월07일

변역출원일자 2002년06월07일

(66) 특허출원번호 PCT/JP2001/10255

(68) 국제출원출원일자 2001년09월06일

(81) 지명국

(87) 국제공개번호 WO 2002/21579

(87) 국제공개일자 2002년03월14일

국제특허 : 중국 일본 대한민국 미국 싱가포르 EP 유럽특허 : 오스트리아 불가리아 스위스 독일 덴마크 스페인 프랑스 영국 그리스 아일랜드 이스라엘 룩셈부르크 모나코 네덜란드 포르투갈 스웨덴 핀란드 사이프러스

(30) 우선권 주장 09/020006.9 2000년09월08일 EP(EP)

(71) 출원인

제닌테크 필립스 일렉트로닉스 컴.부이.

내셔널트랑크, 미안드호판, 그로트보르드스베그 1

(72) 발명자

바르네에리마우르

(74) 대리인

내셔널트랑크, 엠멜-85501아아인드호판, 프롬프. 홀스텐베그

미행로

심사청구 : 없음

(54) 저장 매체상에 저장된 정보 신호를 재생하는 장치

요약

본 발명은 제 1 저장 매체(4)상에 저장된 정보 신호를 재생하기 위한 장치에 관한 것이다. 이 장치는 제 1 저장 매체로부터 정보 신호를 판독하기 위한 판독 유닛(2)과, 정보 신호를 디스클레어 유닛으로 공급하기 위한 출력 유닛(6)과, 정보 신호를 사용자의 액세스할 수 있도록 명령들을 수신하기 위한 사용자 제어가능 입력 유닛(8)을 포함한다. 사용자 제어가능 입력 유닛은 입력의 순서로 제 1 명령을 수신하도록 구성된다. 상기 장치는 상기 제 1 명령에 응답하여 제 2 위치에서 상기 저장 매체로부터 정보 신호를 판독하기 시작하도록 상기 판독 유닛을 제어하기 위한 유닛(10)을 더 포함하며, 상기 제 2 위치에 있는 정보 신호는 상기 제 1 명령의 수신 순간에 판독된 제 1 위치에 있는 정보 신호의 특징들과 유사성뿐 아니라, 또는, 상기 순간 이전에 판독된 정보 신호의 특징들과 유사성을 나타내는 특징들을 가진다.

도면

도 1

색인어

유사성, 히스토그램, 정보 재생, 검색, 판독 유닛

참조부

기술분야

본 발명은 제 1 저장 매체상에 저장된 정보 신호를 재생하기 위한 장치에 관한 것으로, 상기 장치는,

- 제 1 저장 매체로부터 정보 신호를 판독하기 위한 판독 수단과,

- 정보 신호를 디스클레어 유닛에 공급하기 위한 출력 수단과,

- 사용자가 정보 신호를 액세스하는 것을 가능하게 하기 위해 명령들을 수신하기 위한 사용자 제어가능 입력 수단을 포함한다.

또한, 본 발명은 저장 매체상에 저장된 정보 신호를 재생하는 방법, 컴퓨터 프로그램 및 상기 컴퓨터 프로그램들을 실행하는 유체 매체 및 신호에 관한 것이다.

발명기술

통상적으로, 비디오는 기본 비트 레이트 가능, 즉, 채널, 팔리 길이 및 되감기를 사용하여 선택적 방식으로 액세스 및 시청한다.

하드 디스크로 인해 가능한 비디오 레코딩들이 급속하게 출시되고 있다. 이들은 저장된 정보의 양을 한정하지 가지며, 이는 일반적으로 액세스할 수도 있다. 팔리 길이 및 되감기 같은 통상적인 비디오 기능들은 이 액세스를 용이하지 못하며, 시청자들이 비디오 내용을 신속하게 브라우징하는 것을 용이도 못한다.

현재의 경향은 오디오 비디오 정보와 함께, 내용의 설명을 제공하는 것이다(이런 국제 표준 MPEG-7은 멀티미디어 콘텐츠를 위한 기술자(descriptor)들의 표준 세트들 생성하는 것을 목적으로 한다). 이 설명은 가령 시청자들이 수 시간의 기록된 프로그램 내에서 신속하고 효과적으로 검색하는 것을 가능하게 하는 방식으로 용이하게끔 한다. 주된 논점은 사용자-시스템 상호작용이다. 그러나, 사용자 편의하면서, 직관적인 도구를 제공하는 것도 기본 목적을 수행하는 것에 한정되어 있다.

현재까지, 팔리 길이 및 되감기는 페이브 또는 디스크 헤제에서 기록된 비디오들을 브라우징 및 액세스하기 위한, 가장 대중적인, 사용자 용이한 도구들이다. 그러나, 가질 기기들에서 저장할 수 있는 멀티미디어 데이터의 극적인 증가로 인해 이들은 부적합해지고 있다.

발명의 상세한 설명

본 발명의 목적은 그림들이나 슬라이드 쇼 같은 영상들을 가진 수시간의 기록된 프로그램들 또는 데이터베이스들을 같은 비디오들 내에서 신속하고 용이하게 검색하기와 적절한 시청이 용이하면서 직관적인 다른 도구를 제공하는 것이다.

본 발명에 따른 장치는 사용자 제어가능 입력 수단과 원하는 순간에 제 1 영상들 수신하도록 작동되는 것을 목적으로 하며, 이 장치는 정보 신호내에 위치된 제 2 위치에서 상기 저장 매체로부터의 정보 신호를 판독하기 위하여 작동할 수 있는 판독 수단을 제어하기 위한 수단들 더 포함하고, 제 2 위치에서 있는 정보 신호는 상기 순간 이전에 판독된 정보 신호의 일부 또는 상기 제 1 영상들 수신하는 상기 순간에 전역한 제 1 위치에 있는 정보 신호와 유사성을 나타낸다.

본 발명은 하기의 인식에 기반한다. 뉴스 프로그램들, 토크쇼, 가상 예보 같은 다수의 프로그램들은 매우 반복적이 발생한다. 이런 종류의 프로그램들에서, 영상들의 배열은 거의 동일하다. 부가적으로, 동일한 사람이 프로그램에서 연기한다. 뉴스 프로그램에서는 일반적으로 뉴스 앵커들의 변경 통한 연속자가 보인다. 방송자가 존재하는 출현을 알려주므로, 다음 뉴스 앵커들로부터 건너뛰는 것이 가능하다. 이 특성은 사용자들이 비디오 스트림내의 영상으로부터 다음(이전) 유사 영상으로 건너뛰 수 있게 해준다. 본 발명의 발명한 실시예에서, 유사성 기준들(즉, 영상들 사이의 유사성 수준의 평가)은 비디오 스트림부터 추출된 유사 레벨 특성들(즉, 색상, 텍스처, 형상 및 매치들)에 기초하거나, 수평 또는 반지름으로 생성되는 하위 영상 샘플들 같은 보조 정보에 기초할 수 있다. "다음(이전) 유사 영상으로 건너뛰(jump to the next/previous similar image)" 기능은 영상들이 비교되는 방식 및 사용자는 유사성의 기준들이 특화되어 있다. 내용 기반 영상 검색은 멀티미디어 검색 분야에 널리 공지된 기술이다. 본 발명은 그 접근을 사용하며, 사용자들에게 검색한, 직관적이며 사람이 용이한 비디오 데이터 브라우징을 도구를 제공한다.

본 발명의 목적은 비디오 내용에 기초한 비디오 데이터에 대한 액세스 프로세스들을 허용하는 것이다. 통상적 재생 수단, 또는, 예로서, 키-프레임 브라우징 같은 비디오 내용들을 액세스하는 다른 방식을 사용함으로써, 시청자는 현재 영상들 같은 영상(query image)으로서 선택하고, 시스템은 유사 영상들을 검색하기 위해 필요한 작업을 수행한다. 이는 "다음(이전) 유사 영상으로 건너뛰"와 같은 유사성 검색하고, 마지막 영상으로, 이는 비디오 스트림내의 다음 영상으로 건너뛰다. 이 기능은 사용자 편의하게 한다. 즉, 두 개의 비전문 사용자에 구별할 수 있다. 이 방식에서, 프로그램의 관심 부분으로 건너뛰기 위해서, 사용자가 수행하여야 하는 유일한 작업은 비전문을 누르는 것이다. 따라서, 본 발명은 소비자 디지털 비디오 소프트웨어들의 비디오들을 통한 브라우징을 최종 사용자들을 지원하기에 적절히 적합하다.

이들 및 본 발명의 다른 일면들은 도면들을 참조로 하는 세가지 실시예들에 의해 명백 및 명확해질 것이다.

예시적인 도면들을 참조로 본 발명의 실시예들을 보다 상세히 설명한다.

도면의 간단한 설명

- 도 1은 본 발명에 따른 장치의 실시예를 도시하는 도면.
- 도 2는 다음 유사 영상으로 건너뛰 기능을 수행할 수 있도록 해주는 단계들을 도시하는 도면.
- 도 3는 8개의 유사한 영상들의 사적화 항목을 사용하는 후속 절차를 예시하는 도면.
- 도 4는 본 발명에 따른 장치의 다른 실시예를 도시하는 도면.

실시예

도 1은 본 발명에 따른 장치의 실시예를 도시하고 있다. 이 장치는 저장 매체(4)상에 저장된 정보 신호를 판독하기 위한 판독 유닛(2)을 포함한다. 장치는 가변 해상도 포함하는 비디오 레코딩 또는 셋-백 비스 클로부터 공칭된 기능을 가질 수 있다. 정보 신호는 시간기록된 기록 레이어(CD 또는 페이브 같은)나 하드 디스크, 드라이브 같은 기록 장치상에 저장된 TV 신호일 수 있다. 정보 신호는 화면상에 디스플레이 할 수 있는 소정 영상들의 정보일 수 있다. 영혼한 실시예에서, 정보 신호는 영상들의 시퀀스를 포함하는 비디오 신호이다. 그러나, 본 발명은 가변 해상도에 저장된 그림들 또는 슬라이드 쇼들의 일화 같은 소정 영상의 영상들의 집합을 통한 브라우징을 위해 사용될 수 있다. 저장 매체는 하드디스크 드라이브, 예로서, 광 디스크(DVD나 CD 같은)나 고상 메모리 같은 재기록가능 기록 매체의 형태일 수 있다. 그러나, 다양한 정보들 저장하기 위한 영의 다른 적절한 저장 매체가 사용될 수 있다. 저장 매체로부터 판독된 정보 신호는 도 2에서 같은 디스플레이 유닛에 정보들을 공급하기 위해 출력 유닛(6)으로 공급된다. 디스플레이 유닛은 디스플레이

기 장치에 포함될 수 있다.

[illegible][illegible][illegible][illegible]

이러한 배열의 다른 실시예에서, 데이터베이스의 각 열의 모든 정보호출의 위치는 대응한다. '다음으로 건너뛰기' 명령이 수행될 때마다, 배열의 데이터들이 다시 재배열(sorting)된다. 열의 이름으로 순서화한 배열은 순서화된 배열의 위치로 대응하는 열의 이름의 유사성에 의해 결정된다. 한 실시예는 순서화자 기호 유사한 정보 순서의 부호를 건너뛰는 것을 가능하게 한다.

[illegible]

이동은 유사도 점진적 점프(jump to the next similar image) 기능을 다수의 유사한 이미지 사용에
유사도 점진적 점프를 한기와 같다.

- 본소 프로그램에서 몇 개 명칭을 사용하여 다음 본소 아이들용으로 건너뛰는 경우 사용한다. 본소 프로그램

편들은 일반적으로 부제들(subtitles)을 가지고 방송되며, 그래서, 그들에 관한 매우 상세하며 정확한 문자 정보를 추정할 수 있다. 이런 가정에서, 그들에 대하여 '다음(이전) 유사 영상으로 건너뛰기(jump to the next(previous) similar image)' 기능 대신 키워드 기반 검색을 사용하는 것이 합당한 것으로 생각된다. 이것은, 이 두가지는 사용자가 먼저 정보를 얻을 필요 없이 뉴스 중 하나로부터 다음(이전) 것으로 신속하게 건너뛰는 것을 허용한다.

- 고정된 표제를 가진 프로그램내에서 가상 채널, 스포츠뉴스 또는 특정 섹션으로 건너뛰는 것을 허용한다.
- 그 자체의 고정된 크레딧들(credit)나 마질 타이틀을 가지고 있는 프로그램의 시작 또는 끝으로 건너뛰기 위해 사용할 수 있다.
- 비디오 음향의 적합한 음악 프로그램들이 최근 보존되고 있다. '다음(이전) 유사 영상으로 건너뛰기(jump to the next(previous) similar image)' 기능은 하나의 비디오 음향으로 다음 것으로 자동 건너뛰도록 사용할 수 있다.
- 다수의 다국어판과같은 특정 로고와 함께 시작하는 심어진 주제의 특정 섹션들을 가지고 있다. 사용자는 시간 소모적인 놀라움 제거를 허용하지 않고, 다음 주제를 나타내는 로고로 바로 건너뛸 수 있다.

마지막 두 가지의 예들에서와 같이, '다음(이전) 유사 영상으로의 이동 버튼(jump to the next/previous similar image button)'은 하나의 구조(음성인 문자들이 유사한 프레임들에 대응), 즉,

AAAAAAAAbbbbbAAAAccccAAAAddAAAAeeeeeeeeeeAAAAAAAAAAAA...의 구조를 가진 모든 프로그램에서 A 프레임들 사이를 건너뛰기 위해 사용할 수 있다. A 프레임들은 한쌍의 뉴스를 보는 형식에 대응한다. 다국어판들에서, 소문(소문자)이 주제를 제시하는 장면과 대응하며, 이들은 다국어판과 이이들중을 함께 삽입한다. A 프레임에서, 소문(소문자)이 어떤도를 전환하거나, 게스트가 등장하는 TV 쇼에서도 거의 동일한 상황이 이루어진다. 방송에서와 같이, 소문(소문자)이 일반적으로 특정 비디오 음향들을 소개한다. 실제로, 이 구조는 통상적인 방송 TV 프로그램들에서 매우 일반적이다.

비디오를 정렬할때에 구성하고, 단지 다음 장면으로 건너뛰기 버튼을 사용하는 것에 의해 유사한 결과가 달성될 수 있다. 몇몇 예들 예를 보라우징 기능을 다중 장면 또는 카-프레임들에 대한 것 뿐만 아니라, 다음 유사 장면으로 바로 건너뛰는 것을 허용하기 때문에, 보다 통상적인 검색에 비디오에 매우 구성 될수 있는 상이하다. 다음(이전) 유사 영상으로 건너뛰기 가능한 비디오 네비게이션 폭력을 위해 사용할 수 있을 뿐만 아니라 시간을 또는 슬라이드 쇼들을 통한 브라우징을 위해서도 사용할 수 있다.

비디오 프로그램들의 프레임워크들은 프로그램의 시작에 위치되어 있을 때, 또는, 이들이 광고 방송 등의 형태일 때 매우 유용할 수 있다. 집자의 경우는 사용자가 프레임워크부터 장면을 선택할 수 있고, 일단 방송될 때까지 되고 나면, 프로그램내의 유사 영상으로 건너뛸 수 있다. 이 방법으로, 사용자는 그가 원하는 대로 하는 위치에 있는 비디오 프로그램들 보기 시작할 수 있다. 후자의 경우(즉, 뉴스 프로그램의 예들)는, 사용자가 각 카 프로그램내의 관심 부분으로 건너뛰기 위하여 프레임워크부터 선택된 영상을 활용할 수 있다. 이 경우, 프레임워크는 프로그램들을 위한 내용의 네비게이션 간주할 수 있다.

사용자가 일부 영상들을 선택 영상들로서 선택할 수 있게 하는 경우와, 다음(이전) 유사 영상으로 건너뛰기만을 위한 것만 있다. 비디오 스펙트의 관심 부분을 사용하는 대신, 사용자는 선호하는 것들의 서로 서로에서 선택할 수 있다. 이 소위 선호 영상 리스트는 메모리, 하기의 사나리움을 가능하게 한다.

- 사용자가 뉴스 프로그램을 시청하고, 그는 그가 좋아하는 액 텐드가 새로운 비디오 클립을 만드는 것을 발견한다. 뉴스 프로그램은 단지 비디오의 일련 서정(one-minute preview)만을 포함하고 있다. 사용자는 그의 선호 영상 리스트에 비디오의 일부 영상을 저장하고, 그는 새로운 비디오 클립이 전송으로 수 시간의 비디오 음향 건너뛰기(confliction)를 초월한다. 다음번 그는 이미 저장된 영상을 사용함으로써 기록내의 전체 비디오 음향들을 볼 수 있다.

- 사용자가 뉴스 프로그램을 시청하고, 그는 Mc 라렌(McLaren) F1 차량을 가진 하킨(Halkins)이 오주에 매우 큰 사고를 당했다는 것을 알았다. 그는 현재 메모리를 확인하였지만, 그는 시간이 없기 때문에 이를 보기를 원하지 않는다. 이제, 그는 그의 선호 영상 리스트에 있던 뉴스 프로그램으로부터 사고의 영상을 사용하여 저장 사고의 시청으로 건너뛸 수 있다.

다음(이전) 유사 영상으로 건너뛰기 가능한 비디오 서점스의 모든 영상에 다음(이전) 가장 유사한 것을 얻게 것을 필요로 한다. 두 개의 연속적 프레임들은 항상적으로 매우 유사하다. 다음(이전) 유사 영상으로 건너뛸 때, 이들 프레임들은 버려져야만 한다. 한가지 해결 방법은 연속적인 유사 프레임들의 그룹내에서 한 하나의 프레임만을 고려하는 것일 수 있다. 이는 비디오를 몇 번 분할하고, 각 섹스에 대하여 대표적 지점(카-프레임)을 선택하며, 그후, 카-프레임들 사이에서 유사 영상을 검색하는 것과 동가하다.

다음(이전) 유사 영상으로의 건너뛰기 가능한 기술자들이 얻어지는 방식 및 유사성이 측정되는 방식과는 직접적이라는 것을 인지하는 것이 중요하다.

양호한 구현에 있어서, 각 카-프레임으로부터 시작적 기술자가 자동으로 추출한다, 그 시작적 기술자들 사이의 거리와 서점정은 일개성 보다. 낮은 경우와 두 개의 카-프레임들이 유사한 것으로 추정한다. 다음(이전) 유사 영상으로의 건너뛰기 가능한 유사성을 고려할 뿐만 아니라, 프레임들의 상대 위치들도 고려한다. 그 이유는 이것이 한 하나의 다음(이전) 유사 영상들을 검색하여만 하기 때문이다. 두 또는 양쪽 수월한 방법에 수월하는 단계들을 보시하고 있다.

서점정이 다음 유사 영상으로 건너뛰기 버튼을 누를 때, 서점들은 이리저리 상들의 기술자들을 검색하며, 두가지 클러스터링을 수행한다. 첫 번째는, 클러스터링은 클러스터링의 시작적 기술자들 사이의(선형)는 카-프레임들의 기술자들과 비교한다. 그 기술자들에 고정된 일개성 보다 큰 클러스터링으로부터의 거리들 간의 카-프레임들은 버려진다. 두 번째 클러스터링은 나머지 카-프레임들을 특성 공간(feature space)내에서 그 클러스터들의 거리에 따라 적어도 두 개의 클러스터들을 분할하는 것으로 이루어진다. 두 개의 클러스터들은 그 유사성에 따라 영상들을 재배열하고, 그 클러스터들의 거리들 사이의 연속적 편차를 고려

임으로써 얻어진다. 이들 편차들 중 하나가 소정 임계값을 초과할 때, 모든 연속적인 영상들이 하나의 다른 클러스터내로 집어넣어진다. 집어넣은 가장 가까운 영상들의 클러스터가 시간적 순서(chronological order)에 따라 재배열되며, 실제 첫 번째 프레임이 다른 유사도를 제공하는 것이다.

하기에서 설명될 수 있는 시각적 기술자들에 관한 모두 세부 사항들을 설명한다.

현재 패턴 형성 및 영상 이해 기술들은 여전히 의미적 용어들의 시각적 내용을 해석하는 목적과는 괴리되어 있다. 따라서, 포우-레일 시각적 특성들에 의존할 필요가 있다. 색상, 텍스처, 영상 및 움직임이 가장 일반적으로 사용되는 시각적 시각 특징들이다. 색상 정보는 영상 크기, 방향성 및 폐쇄(occlusion)에 대해 보다 강인성을 가진다. 텍스처 기술자들은 관심 텍스처 패턴들을 분류하는데 강해하지만, 그러나, 이들은 시간 및 공간 둘둘 배정된 데이터에는 효과적이지 못하다. 또한, 영상 기술자들이 사용할 수 있다. 이들 기술자들이 영상 세그먼테이션을 필요로하고, 이 영상 세그먼테이션이 많은 연산작을 필요로하기 때문에, 이들은 현재 소비적 전자 제품들에는 너무 고가이다. 부가적으로, 현재 이점가능한 영상 세그먼테이션 기술들은 완전한 실제계(real-world) 영상들에 대해서는 충분히 강력하지 못하다.

언간의 색상 인지는 매우 프로세싱이다. 시각적 데이터 및 관련 표현들을 허용할 때, 다수의 단순화 가정들이 이루어진다. 색상 특성들은 항상 레이블에서 제거되며, 이는 색상의 인지가 주변 색상들에 의해 영향을 받지 않는 것을 의미한다. 부가적으로, 주변성, 연속 거리 및 다스플레이 중심 같은 세계 조공들은 고려되지 않는다. 시각적 의미들의 몇 번째 경우가 주어진다. 시각적 의미들(1)은 전체 영상 또는 영상 공간(CS)내의 색상값들의 세트에서 표현되는 영상의 소량의 임의 영상 영역(세시각적 또는 불규칙적)이다.

하기의 문단들은 시각적 시각적 특성들의 광범한 표현들을 인코딩하는 기술자들의 세트를 제공한다. 추출하기 및 일정한 유사성 측정 기준들도 제공한다.

원래 히스토그램들은 시각적 의미들의 매우 광범한 색상 특성을 설명하는 널리 공지된 방식이다. 이는 색상 채널들에 관한 하나의 분포 또는 세 개의 독립적인 색상 분포들로서 표현될 수 있다. 원래 히스토그램들은 n 개의 별개의 색상들이 존재하도록 아산된. 색상 공간(CS)내의 주어진 시각적 의미들(1)을 위해 정의된다. 원래 히스토그램(H(1))은 벡터 $\{r, g, b, \dots, k\}$ 이며, 여기서, 각 요소(k)는 시각적 의미들(1)내의 색상(C),의 최소들의 비율을 포함한다.

원래 히스토그램들은 색상 내용의 보다 효과적인 표현이다. 한가지 긍정적 특성은 그 연산이 효율적이라는 것이다. 부가적으로, 원래 히스토그램들은 카메라 회전, 질, 해상도의 변경 및 부분적 폐쇄에 의해 발생하는 변화들에 민감한 영향을 받지 않는다. 그러나, 이들은 관 조공들에 민감하며, 색상 공간 위치화치에 의해 색상 내용의 표현에 문제가 있을 수 있다. 원자화는 반드시 시각적으로 별개의 색상들이 동일 본(bn)내에 존재하지 않을 정도로 충분히 미세하여야만 한다. 이 고려 사항은 하기의 장들에 제시되는 모든 히스토그램 기반 기술자들에 적용될 수 있다.

원래 히스토그램 추출은 시각적 의미들내의 각 최소값에 대하여 양자화된 값을 얻고있다. 히스토그램내의 대응 값을 증분시키는 것에 의해 수행된다. 그후, 벡터내의 최소들의 수는 시각적 의미들의 크기에 따라 정규화되어야만 한다. 이 마지막 단계는 동일한 차수의 시각적 의미들들을 취급하는 경우에 적용될 수 있다. 추출 절차는 선택적 시간을 필요로한다.

원래 히스토그램들을 사용하여 색상적 유사성을 결정하기 위해 상이한 거리 척도들이 사용될 수 있다. 이들은 연산적 적합성과 효율성 양자 모두에 관하여 상이한 성격 상충들을 초래한다. 색상 공간 및 색상 차이의 선형과 비선형, 유사성 측정 기법의 선택은 시각적 질의 기술의 구현에 결정적인 특성이다.

원래 히스토그램들에 대한 동일한 측정에 항상적으로 사용되는 세가지 유사성 척도들은 L_1 거리, 유클리드 또는 L_2 거리 및 기타 거리(quadratic distance)이다. $H(1)$ 및 $H(1)$ 를 각각 줄로 및 목표 히스토그램이라 하면, 아래, L_1 은 하기와 같이 정의된다.

$$D_1 = \sum_{i=1}^n |H(I_q) - H(I_t)| \quad (2.1)$$

유클리드 거리 또는 L_2 거리는 하기와 같이 정의된다.

$$D_2 = \sqrt{\sum_{i=1}^n (H(I_q) - H(I_t))^2} \quad (2.2)$$

이들 정의들에서, 색상 벡터들을 가리키는 편차들은 공통하게 가중된다. L_1 및 L_2 거리들 양자 모두는 유사하지만 유클리드 양자 히스토그램 요소들을 비교하지 않는다. 그래서, 이들은 적색 영상은 녹색 영상과 마찬가지로 동일한 동등하게 처리된다. 그러나, 연산내의 히스토그램 요소 유사성의 척도들을 사용함으로써, 히스토그램 행렬을 형성시키는 것이 가능하다.

마지 거리 척도는 이 논점에 주안점을 두고 있으며, 하기와 같이 정의된다.

$$D_3 = (H(I_q) - H(I_t))^T (H(I_q) - H(I_t)) \quad (2.3)$$

여기서, $H(I_q)$ 는 $n \times 1$ 인 벡터 i 를 가지는 색상들의 시각적 유사성을 나타낸다. 이 행렬들은 모든 히스토그램 요소들을 비교하고, 요소의 거리를 행간과 기중 연산들에 의해 요소들 상동한 거리들을 가중한다. 상동 관련(α_{ij})을 위한 특별한 값은 $\alpha_{ii} = 1$ 에 의해 주어지며, 여기서, α_{ij} 는 색상 공간상동 두 색

상들 사이의 최대 거리에 관하여 정규화된 유클리드 1차 1의 역상을 사이의 거리이다. 이차 거리는 모든 요소들 사이의 상호 유사성을 연산하기 때문에, L_1 및 L_2 보다 연산적으로 보다 많은 비용을 소모한다.

MPEG-2 표준에 채택된 압축 알고리즘이 사용될 때, 단지 부분적 디코딩만으로도, 지각적 DC-성상들이면 지각되는 프레임들의 축소 재생제곱 비전들(0.4배 작음)을 비디오 스트림으로부터 추출하는 것이 가능하다. 이들은 전체 크기의 프레임의 8x8 블록들의 양방향 1차인 코사인 변환의 DC 계수들만을 고려함으로써 얻어진다. DC 성분들이 보다 작게 축소 재생제곱되거나, 프레임들의 블록 해상 비전들이기 때문에, 이들은 동일한 내용물 포함하는 것으로 가늠할 수 있다. 유사한 기-프레임들을 검색하는 목적을 위해서, 디지털 비디오 스트림내에서 지각적, 1-프레임들을 위해, 이용가능한 DC 성분들로부터 직접적으로 지각적 기술자들을 추출하는 것이 가능하다. MPEG-2 스트림에서, 1-프레임은 기-프레임으로서 간주된다. 그 구현을 단순화하기 위해서, 후술되는 특정한 역상 공간내의 최소값들의 아래에서 압축으로서 참조하도록 설계된다. 또한, 그들을 정면 변경 알고리즘과 통합하고, MPEG-2 디코딩을 최소 필요치로 제한하는 기술자들의 연산을 수행하는 것도 가능하다.

컬러 히스토그램 기술자(colour histogram descriptor)는 Y, C_b 및 C_r 역상 공간들 양자 모두에 사용될 수 있다. Y, C_b 역상 공간은 MPEG-2 표준에 사용되는 것임이므로, 따라서, 비디오 스트림으로부터 직접적으로 추출된 역상 정보가 추가적인 변환을 필요로하지 않을 때 적합하다. 더욱이, 엄격히 자각적으로 동일하지 않은 경우에도, 이 관점으로부터, 사용자 인터페이스내에 기-프레임들을 디스플레이하기 위해 사용되는 RGB 역상 공간보다 선호한다.

다름에, 사용될 수 있는 Y, C_b 역상 공간의 세가지 삼이한 역상 양자화율이 정역된다. 다른 양자화율도 마찬가지로 적합할 수 있다는 것을 인지하여야만 한다.

- Y, C_b 및 C_r 역상 채널들이 각각 16, 4 및 4 레벨들로 선형적으로 양자화된다. 결과적인 기술자는 256 레벨들의 단일 가변 히스토그램들로서 나타난다.

- Y, C_b 및 C_r 역상 채널들이 각각 16, 8 및 8 레벨들로 선형적으로 양자화된다. 결과적인 기술자는 1024 레벨들의 단일 가변 히스토그램들로서 나타난다.

- Y, C_b 및 C_r 역상 채널들이 각각 16, 16 및 16 레벨들로 선형적으로 양자화된다. 결과적인 기술자는 4096 레벨들의 단일 가변 히스토그램들로서 나타난다.

HSV 역상 공간도 마찬가지로 적합하며, 그 이유는 실질적으로 자각적으로 동일하고, 따라서, 엄격적이고 완전한 역상들의 집합이 적절한 양자화를 정의함으로써 얻어질 수 있기 때문이다. RGB로부터 HSV로의 변환은 하기의 식들[36]을 통해 달성된다.

$$v = \max(r, g, b)$$

$$s = \frac{v - \min(r, g, b)}{v}$$

$$h = \begin{cases} r = \max(r, g, b) \text{이고 } g = \min(r, g, b) \text{인 경우 } 5 + b' \\ r = \max(r, g, b) \text{이고 } g \neq \min(r, g, b) \text{인 경우 } 1 - g' \\ g = \max(r, g, b) \text{이고 } b = \min(r, g, b) \text{인 경우 } 1 + r' \\ g = \max(r, g, b) \text{이고 } b \neq \min(r, g, b) \text{인 경우 } 3 - b' \\ b = \max(r, g, b) \text{이고 } r = \min(r, g, b) \text{인 경우 } 3 + g' \\ \text{기타 } 5 - r' \end{cases}$$

여기서, (r, g, b) 는 RGB 공간의 자질이고, (h, s, v) 는 HSV 공간내의 대응 자질이며, (r', g', b') 는 $\max(r, g, b) = \min(r, g, b)$ 일 때 이거와 같이 정의된다.

$$r' = \frac{v - r}{v - \min(r, g, b)}$$

$$g' = \frac{v - g}{v - \min(r, g, b)}$$

$$b' = \frac{v - b}{v - \min(r, g, b)}$$

$r, g, b \in [0 \dots 1]$ 인 경우에서, 이 변환은 $h, s, v \in [0 \dots 1]$ 을 재공간화한다.

이전 컬러 히스토그램 연산하기 위해 필요한 역상 공간 양자화는 168 역상들의 축소 세트[5, 7, 8, 19,

23]을 생성하도록 설계된다. 빛샘이 자각적으로 보다 현저한 특성이 것으로 생각하면, 이를 위해 가장 미세한 입자화가 사용된다. 원형형 HSV 색상 공간의 빛샘 원은 20°의 단계를로 분할된다. 이 방식으로, 심원적 및 노반, 마르타 및 사빈에 따라 세 개의 서로 다른 단계를를 가지는 형태로 분할된다. 채도 및 밝은 각이 세 개의 레벨들로 입자화되며, 마르타, 사빈, 채도 및 밝은 각이 세 개의 레벨들로 입자화된다. 16 빛샘들, 5 채도, 3 밝은 각과 4 부가적 그레임을 사용함으로써, 10672개의 빛샘들과 그레임들이 생성된다(18×5×3×4×16).

세 개의 YC₀ 및 YC₁ 및 HSV 색상 공간을 고려함으로써 얻어진 세 개의 컬러 히스토그램들이 전체 크기 입상들과 DC-입상들 양자 모두로부터 추출된다. 따라서, 각 카-프레임은 0개의 상이한 컬러 히스토그램들이 얻어진다.

상이한 입상들의 컬러 히스토그램들은 L, 및 율리드 거리들을 사용함으로써 비교할 수 있다. YC₀ 색상 공간내의 세 개의 입상들을 사이에서, DC-입상들로부터 추출되고, L, 거리와 비교된 256 번째 컬러 히스토그램들을 사용하여 최상의 결과를 얻어진다. 따라서, L, 거리는 보다 고가의 율리드 거리 보다 양호하게 수행되는 것으로 판명되었다. 또한, 전체 크기 프레임들이 아닌 DC-입상들로부터 히스토그램들을 추출하는 것이 이 기술자의 컬러 성능들을 악화시키지 않는 것으로 판명되었다.

DC-입상들로부터 추출된 HSV 색상 공간내의 108 번째 컬러 히스토그램은 YC₀ 색상 공간의 것 보다 양호한 것으로 판명되었다. 이 결과는 HSV 색상 공간이 충분히 자각적으로 균일하며, 106개의 빛샘들과 색상들이 YC₀ 색상 공간의 256 빛샘과 보다 양호한 색상 범위를 제공한다는 사실로 인한 것이다.

비록, 컬러 히스토그램에 의해 제공되는 것 같은 포괄적 색상 설명은 신축성있게 구분할 수 있지만, 공간적 심보의 결여는 시각적 의미들의 비교시 너무 많은 오류들을 줄 수 있다. 전체 흐름 및 정렬을 향상시키기 위해서, 색상 특성과 공간적 관계까지 모두가 사용될 수 있다. 컬러 그레임 히스토그램과 함께의 히스토그램 기반 기술자들은 공간적 심보를 비잔가지로 위험으로부터 포괄적 색상 특성을 지각한 것으로 향상된다.

주어진 컬러 히스토그램에 의해 표현된 전체적 색상 특성을 확장하기 위하여, 자연적 접근 방법은 시각적 의미들을 세로로 분할하고, 세로 블록들 각각으로부터 색상 특성을 추출하는 것이다. 컬러 그레임 히스토그램은 이 접근 방법을 따르며, 이는 세 개의 컬러 히스토그램들로 구성될 수 있다. 이 경우에서, 시각적 의미들은 3×3 정사각형 그레임을 사용하여 아홉 개의 영역들로 분할된다. 각 영역으로부터 전체의 컬러 히스토그램이 얻어진다. 일반적으로, 컬러 히스토그램은 전체 시각적 의미들의 컬러 히스토그램이다. 세로 블록 분할은 종종 블록들의 그리드들로부터 계산될 수 있다. 이 접근 방법에서, 기술자는 작은 영역 분할에 비교해 좀 더 쉽다.

컬러 그레임 히스토그램 추출 결여는 컬러 히스토그램에서도 실질적으로 동일하다. 단지 차이점은 시각적 의미들의 컬러 히스토그램들의 공간적 위치에 따라서도 히스토그램 요소들이 증분한다는 것이다.

컬러 그레임 히스토그램이 컬러 히스토그램들로 구성되지 않기 때문에, 히스토그램들을 비교하기 위해 사용되는 동일 거리 메트릭들이 세로 블록 히스토그램들에 사용될 수 있다. 두 개의 컬러 그레임 히스토그램을 사이에서 거리는 세로 블록 거리들을 줄 합계이다. 부가적으로, 영상 내의 그 위치에 따라 세로 블록 히스토그램들 사이에서 거리를 가중할 수 있다. 채도, 밝은 각과 다른 것 보다 밝은 밝은 가중할 수 있으며, 그 이유는 대부분의 시각적 의미들에서, 중앙 영역이 관측자의 주의를 끌고, 그래서, 인간들은 유사성 측정의 중앙에 보다 많은 비중을 두기 때문이다.

컬러 그레임 히스토그램 기술자는 3×3 정사각형 그레임을 사용하여 전체 크기의 카-프레임들용 0개의 영역들로 분할하고, 각 세로 블록에 대해 64 번째 컬러 히스토그램을 얻어냄으로써 구성된다. 부가적으로, 전체 영상에 대하여 다른 64번째 히스토그램이 얻어진다, 따라서 기술자는 10개의 히스토그램들로 구성된다.

각 히스토그램은 YC₀ 색상 공간내에서 얻어진다. Y, C, 및 C, 색상 채널들은 각각 하나씩의 4 채널들로 생성함으로써 입사된다. YC₀ 색상 공간은 MPEG-2 스타일내의 색상 정보가 이 요트로써 이용가능한 해 사용되는 것이 적절하다.

다른 입상들의 컬러 그레임 히스토그램들을 비교하기 위해 사용되는 거리들은 대응 세로 영역 히스토그램들 사이에서 율리드 거리들의 또는 L, 거리들의 합이다. 부가적으로, 영상내의 그 위치에 따라서 세로 블록 히스토그램들 사이에서 거리를 가중한다. 중앙 블록 거리는 나머지들 보다 2 배지 10배 가중치된다.

실제 테스트들은 동일 영상들에 대하여, 컬러 그레임 히스토그램이 보다 양호한 결과들을 제공한다. 비균적으로 다른 컬러의 것 보다 양호하게 수행되지 않는다는 것을 보여준다. 거리에 연산에 상이한 가중치들을 사용하는 것에 대해서도, 전체 흐름은 그 추출, 비교 및 지각의 부가적인 비용을 고려할 때 충분한 개선되지 않는다.

컬러 구조 히스토그램들(또한, 문헌에서는 블로브 히스토그램들(blob histograms)이라 명명되는)은 다수의 히스토그램들 구성되어 있는 구조 요소들을 사용하여 시각적 의미들의 지각적 색상 구조를 표현한다. 컬러와 컬러 히스토그램들은 특정 색상들 가진 단일 히스토그램들의 상대적 빈도수를 측정한다. 색상 구조 히스토그램들은 일련의 특정 색상들 가진 히스토그램을 포함하는 구조 요소들의 상대적 빈도수를 연구하기 때문에 명명되는 것된다. 이들은 컴퓨터 컬러 히스토그램들로부터 블로브 특성을 계산하며, 공간적 심보를 이해함으로써, 그 블로브(blur)를(haziness)를 측정해 줄 수 있다.

그들이 세 개의 컬러의 색상들이 되도록, 컬러 색상 공간(CCS)을 고려하면, 시각적 의미들(1)을 위한 컬러 구조 히스토그램은 두개의 컬러와 관련될 수 있다.

정의 2' 색상 구조 히스토그램(1)은 벡터 $\{h_1, h_2, \dots, h_n\}$ 이며, 여기서 각 요소 $\{h_i\}$ 는 색상 $\{C_i\}$ 와 하위 이성의 화소들을 포함하는 시각적 이미지(1)내의 구조 요소들의 수를 표현한다.

구조 요소의 공간적 범위는 시각적 이미지(1)내의 크기와 의존하지만, 그러나, 동시에 구조 요소와 시각적 이미지(1)의 서브샘플링으로 인해 구조 요소들의 선택은 수에 영향을 미친다. 8x8 픽셀로 이루어진 구조 요소 내에서 다수의 54 샘플들을 선택하는 경우, 이 패턴에서 두 샘플들 사이의 거리는 시각적 이미지의 크기를 증가시킬 때 따라 증가한다. 시각적 이미지들이 고정된 기본 크기로 크기 재설정되는 경우, 동일한 8x8 구조 요소가 사용될 수 있고, 한편 서브샘플링 인자 및 구조 요소 폭 및 높이는 하기의 값에 설정될 수 있다. 예를 들어, 시각적 이미지(1)내의 구조 요소의 공간적 범위가 하위, 공간적 범위는 8x8이다. K를 색상의 서브샘플링 인자로 하면, $K=1, 2, 4, 6, 10, \dots$ 이고, 여기서, $K=1$ 은 이미지(1)의 서브샘플링도 있는 것을 의미하며, K=2는 수평적 및 수직적으로 2만큼 서브샘플링된다는 것을 의미하는 등등이다. K와 L은 하기의 값에 설정된다.

$$p = \max\{0, \text{round}(0.5 \cdot \log_2(\frac{K}{L} \cdot \frac{H}{W})) - 8\}$$

$K=2p$

$L=8K$

여기서, 폭 및 높이는 시각적 이미지(1)에 관련 것이다. $p=0$ 인 경우에, $p=0$ 로 생각한다.

결과 구조 히스토그램은 시각적 이미지의 모든 위치들을 방문하고, 각 위치에 대해 동일한 구조 요소내에 포함된 모든 화소들의 색상들을 검색하고, 대응 방울을 방문함으로써 연산된다. 히스토그램 방문은 각자의 종표시 구조 요소들의 수에 의해 정규화될 수 있다. 도 3은 8개의 상이한 색상들의 시각적 이미지들(1)을 방문 후의 결과를 보여주고 있다.

4x4 화소들의 시각적 이미지(1) 구조 요소(2)가 슬라이딩 윈도우에서 시각적 이미지(1)를 통과한다. 색상의 위치(도면)는 단지 시각적 이미지의 앞부분이 도식화(1)에서, 구조 요소는 색상 C_0 와 4 화소들의 4 화소들 및 색상 C_1 와 8 화소들을 포함한다. 이때, 열들 $\{C_0, C_1\}$ 의 방문은 순번된다. 그래서, 이 경우, 구조 요소(structuring element)는 구조 요소 영역내에 존재하는 각 색상에 대하여 한번씩 세로로 계수된다(counted).

색상 구조 히스토그램들, 색상 상관도들(colour correlations), 색상 자기상관도들(colour autocorrelations), 색상 동적 벡터들(colour coherent vectors) 및, 점진적 히스토그램들(accumulated histograms) 기반 기술자들이기 때문에, 종래의 컬러 히스토그램들을 위해 제공되는 동일한 유사성 측정 기술들이 다른 다른 다른 시각적 기술자들을 비교할 때에도 적용될 수 있다. 서로 다른 특성 공간들내의 거리값들은 물론 비교될 수 있다.

색상 상관도는 색상 쌍들에 의해 만족된 이미지이며, 여기서, $\langle i, j \rangle$ 에 대한 k 번째 엔트리는 색상 C_i 와 화소로부터 거리 k에 있는 색상 C_j 의 화소들 발견 가능성들을 측정한다. 색상 상관도들은 색상의 공간적 상관 관계가 거리에 따라 변화하는 방식을 나타낸다.

주어진 시각적 이미지(1)와 n 개의 별개의 색상들이 존재하도록 미분화된 색상 공간(C_S)이 주어지고, i, j 가 n개의 색상간의 색상을 나타내는 것이라 한다. 따라서, 표시된 $p \in P_i$ 는 $P \in P_i$, $i(p)=n$ 과 같은 의미이다. 화소들 사이의 거리는 화소들 $p_1(x_1, y_1), p_2(x_2, y_2)$ 에 대하여 L_2 -norm으로 측정되며, 하기의 값이 정의된다.

$$\|p_1 - p_2\| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$$

상기 세트 $\{1, 2, \dots, n\}$ 을 $[n]$ 으로 나타낸다. 이러한 표기법, 결과 히스토그램 $H(i) = \forall i \in [n]$ 에 대하여 하기의 값에 정의된다.

$$h_n(j) = \Pr\{p \in I_{c_j}\}$$

시각적 이미지(1)내의 일련의 화소들에 대하여, $h_n(i)$ 는 화소의 색상이 c_n 일 가능성을 제공한다. 거리 $d \in [n]$ 을 우선적으로(a priori) 고정되게 한다. 이때, i의 색상 상관도는 $\forall i, j \in [n], k \in [d]$ 에 대하여 하기의 값에 정의된다.

$$\gamma_{c_i, c_j}^{(k)} = \Pr_{p_1 \in I_{c_i}, p_2 \in I_{c_j}} \{ \|p_1 - p_2\| = k \}$$

영상내의 색상 c_i 와 경로의 화소가 주어진다면, $\gamma_{c_i, c_j}^{(k)}$ 는 주어진 화소로부터 거리 k에 있는 화소가 색상 c_j 로 이루어져 있을 가능성을 제공한다.

시각적 이미지(1)내의 상이한 색상들의 수가 있을 때, 색상 상관도와 공간적 및 시간적 연산적 복잡성은 증가한다. 즉, 복잡성들은 감소한다. 이는 동일한 색상들 사이의 상관관계들을 고려함으로써 부분적으로 해결될 수 있다. 이 색상 상관도의 측수는 색상 자기상관도를 지칭된다. i의 자기상관도는 동일한 색상들의

이의 공간적 상관관계만을 포착하며, 이는 허가와 같이 정적된다.

$$\alpha_c^{(k)}(f) = \gamma_c^{(k)}(f)$$

시각적 의미에 관계없이 색상 분포에 관한 공간적 정보를 포함시킴으로써, 색상 상관도들과 자기상관도들은 특히 유사한 색상을 가지지만 컬러 레이어마다 상이한 시각적 이미지를 제공하는 채 컬러 히스토그램들 보다 많은 정보를 제공한다.

시각적 의미들(1)의 색상 상관도를 연산하기 위한 naive 알고리즘(naive algorithm)은 색상 c_i 의 각각의 p_i 에 대한 것들 고려하며, 각각의 $k \in [d]$ 에 대하여, $\|p_i - p_k\| = k$ 로써 색상 c_i 의 모든 $p_k \in I$ 를 계수(count)한다. 대안적으로, 이는 $O(d^3)$ 시간을 취하며, 여기서 d 는 1의 화소값들의 총수이다. 이 고가의 연산을 제거하기 위하여, International Journal of Computer Vision Vol.35, 1999, pp 2845-2888의 주어진 색상 인덱싱 및 어플리케이션(Spatial Colour Indexing and Application)이란 논문에 동적 프로그래밍에 기반한 효과적인 알고리즘이 서술되어 있다. 소위 연산 시간은 $O(d^2)$ 로 감소한다.

같은 효율에 관련하여, L_1 거리들 가진 색상 자기상관도와 종래의 히스토그램보다 상호하게 실행되는 것으로 증명되었다. 그 때문에 불구하고, 그 연산, 메모 및 저장의 부가적인 비효율들을 고려하면, 점들의 가중치는 같고도 효과적인 컬러 히스토그램이다.

널리 공지된 색상 일치 벡터들은 기본적으로 색상 분포에 관한 일부 공간적 정보를 포함하도록 확장된 컬러 히스토그램들이다. 색상의 중심은 그 색상의 화소들에 시각적 의미들(1)의 내림 유사 채색 영역들의 구성원에 의해 정도로서 정의된다. 시각적 의미들(1)을 고려하면, 컬러 히스토그램(H(1))의 주어진 색상 바깥내의 각 화소는 대략 유사 채색 영역의 일부인지 아닌지에 기초하여, 중심성 또는 비중심성 중 어느 한 쪽으로 분류된다. 색상 일치 벡터(COV)는 각 색상을 가진 원상 내 비중심 화소들의 벡터를 저장한다. 1번째 이산된 색상의 합성 화소들의 벡터(α_i)과, 비중심 화소들의 벡터(β_i)를 조합하면, 색상 일치 벡터는 각 이산된 색상에 대해 하나의 벡터들의 쌍들로 표현될 수 있다.

$$\langle \alpha_1, \beta_1 \rangle, \dots, \langle \alpha_n, \beta_n \rangle$$

상동들(α_i 및 β_i)에 대하여 동일 의미를 유지시킴으로써 종래의 컬러 히스토그램들 벡터에 의해 허가와 같이 표현될 수 있다.

$$\langle \alpha_1, \beta_1 \rangle, \dots, \langle \alpha_n, \beta_n \rangle$$

색상 일치 벡터들(Colour Coherence Vectors)(CCV)은 일 시각적 의미들의 합성 화소들에 다른 것들내의 비동성 화소들과 대칭되는 것을 방지한다. 동일 화소들을 비동성 화소들로부터 분리시킴으로써, CCV는 컬러 히스토그램들 보다 세밀한 분별을 제공한다.

색상 일치 벡터를 추출하는 첫 번째 단계에서, 시각적 의미들(1)은 화소값들을 작은 지역적 이웃(통상적으로 8 픽셀 가진 인접 화소들)내의 평균값으로 교체함으로써 다소 희박하게 된다. 이는 이웃 화소들 사이의 다른 변화들을 소거한다. n 개의 별개의 색상들의 이산된 색상 공간이 사용되는 것이 타당하다.

다음 단계는 주어진 색상 바깥내의 화소들을 중심성 또는 비중심성으로 분류하는 것이다. 중심성 화소는 동일 색상의 화소들의 큰 그룹의 일부이고, 비중심성 화소는 아니다. 인접된 화소들만을 연산함으로써 화소 그룹들을 결정한다.

정의 3 : 연결된 컴포넌트 C는 영역의 두 개의 화소들 $a, b \in C$ 에 대하여 a 와 b 사이의 C에 하나의 경로가 있도록 화소들의 확대화된 세트이다.

정의 4 : C내의 경로는 각 색상 $p \in C$ 에 일의 두개의 순차 화소들 p_i, p_{i+1} 이 서로 인접하는 화소들의 시퀀스 $p_0, p_1, p_2, \dots, p_n = p$ 이다. 하나의 화소가 n 개의 다른 화소들 중 여러 개의 가장 근접한 이웃들 사이에 있는 경우 두 개의 화소들만 인접한 것으로 간주한다.

주어진 색상 바깥내의 연결된 컴포넌트들만이 연산된다는 것을 주목하라. 연결된 컴포넌트들만 연산이 종료되었을 때, 각 화소는 정확하게 하나의 연결된 컴포넌트에 속한다. 화소들을 연결된 컴포넌트의 화소들내의 크기(여기서, 중심성 또는 비중심성 중 어느 한 쪽으로 분류한다. 그 연결된 컴포넌트의 크기가 고정된 값(τ)을 초과하는 경우 화소는 중심성이며, 그렇지 않으면, 화소는 비중심성이다. τ 는 대개 시각적 의미와 고가의 1픽셀 설정된다.

L_1 거리에 대해, 1픽셀 해상도(하나의 332 픽셀 벡터를 제공)를 양자화된 HSV 색상 공간내의 색상 일치점들 검색 효율성, 검색적 비용 및 응답 시간에 관하여 시험된 화상의 시각적 유사성을 것으로 판명되었다.

종래의 컬러 히스토그램들과 유사하게, 색상 일치 벡터들은 색 조건들의 변화에 민감할 수 있다. 광 역학적 색상 특성들을 유지하기 위한 방식은 HSV 색상 기술내의 밝기와 채도 컴포넌트들만을 사용하여나, 그 한층 확대 3D의 색상 공간의 적색, 녹색 및 청색을 정규화하는 것이 될 수 있다.

색 분포에 대한 속한자들에게 널리 공지되어 있는, 조인트 히스토그램들은 색상 일치 일치 벡터들과 컬러 히스토그램들의 쌍방이다. 색상 일치성 뿐만아니라, 지역 화소 특성들의 세트들 고려함으로써, 이들은 다차원 히스토그램들로서 분류될 수 있다. 조인트 히스토그램내의 각 임의의는 특정값들의 특정 조합에 의해 기술된 영상내에 다수의 화소들을 포함한다. 보다 정확히히는, i 번째 특성이 n_1 가능값들을 가지는 k 색

상들의 세트가 주어질 때, 조인된 히스토그램은 k 차원 벡터이며, 그래서, 조인된 히스토그램의 각 연립 리기 특성값들의 k 개의 집합에 대해 설명되는 시각적 이미지들의 화소들의 배열을 포함한다. 따라서, 조인된 히스토그램의 크기는 $n \times n \times n_1 \times n_2$ 이고, 각 특징의 값들의 가능한 조합들의 수이다. 그러나 히스토그램이 화소 색상들의 일부를만 포함하기 때문에, 조인된 히스토그램은 몇몇의 화소 특징들의 조인된 배열을 포함한다. 색상 중심 벡터는 색상들과 색상 응답변환을 특징들로서 사용하는 조인된 히스토그램으로 될 수 있다. 컬러 그러나 히스토그램들은 특정 서브 영역에 속하는 위치와 색상들을 특징들로서 사용하는 조인된 히스토그램으로 될 수 있다. 예를 들어, 합도(contrast)와 엣지(edges)는 화소들로부터는 작은 이미지들의 화소들의 위치와 배열을, 텍스처도니스(textures)와 텍스처도니스(texture)는 그 강도가 화소들보다 많이 상이한 연속 이웃 화소들의 수를, 구배 크기(gradient magnitude) (구배 크기는 최대 변화의 방향으로 강도가 얼마나 신속하게 변화하는지의 척도인), 프랭크(franks) (색소 p 의 등등은 그 강도가 p 의 강도 보다 작은 지리적 이웃 내의 화소들의 수로서 상의될) 등의 일부 부가적인 특징들을 사용함으로써, 조인된 히스토그램들은 색상 중심 벡터들 보다 더 많은 판별을 제공한다.

시각적 미미함으로부터 조인된 히스토그램들을 추출하기 위한 절차는 시각적 내용을 특징시키기 위해 선택된 특징들에 의존한다. 선택적 사건에서 효과적으로 연산될 수 있는 특성적인 특징들이 선택된다.

조인된 히스토그램들로서 볼 수 있는 컬러 그러나 히스토그램들과 컬러 응답성 벡터들과는 별도로, 색상, 색상 응답성 및 평균 텍스처 복잡성(average texture complexity)을 특징들로서 사용하는 조인된 히스토그램이 사용된다. 평균 텍스처 복잡성은 $N \times N = 2$ 바이트로 스토리징하는데 적합한 일련된 도해된 점들의 집합을 의미하며, 우정된다. 각 블록의 이전 요소와 변위된 $N/4$ 개수들은 그림이 그림은 원형일 때 있는 경우 최대 값으로 설정되고, 그림이 일정한 배열에 있는 경우 0으로 설정된다. 블록은 이미지의 화소들의 수가 다른 사건들만큼 길게될 때 있는 경우 '복합(cosine)'으로서 평가된다. 색상 및 색상 응답성이 부가적으로, 카-프래임의 각 화소는 그것이 속하는 블록의 텍스처 복잡성에 따라 두 개의 블록으로 분류된다.

조인된 히스토그램의 평균 컬러 응답성은 색상 응답성 벡터들을 채용함으로써 얻어진 것들과 비교할만 하다. 부가적인 텍스처 특징은 반복적으로 향상시키지만, 상기 특징들을 구현하기 위한 비용은 상대적으로 높을 수 있다.

1-프래임들과 카-프래임 컬러 유효성을 향상시키기 위해 이용된다. 카-프래임들이 특정 기간들에 따라 선택되지 않는 경우에, 이때, 이웃하는 1-프래임들의 시각적 기술자들을 사용하는 것이 다들 (미리) 유사 영속도 근대적 기술의 강력 효과를 향상시킬 수 있다. 각 카-프래임을 위한 단 하나의 시각적 기술자를 고려하는 대신, 상기 카-프래임에 근접한 1-프래임들과 그들의 모든 시각적 기술자들이 연산되고, 부가적으로, 그 크기가 화소들 이상 큰만큼 기술자들이 모여 합성된다.

이 기술의 보다 상세한 설명이 이따간다. 각 섷(shot)에 대하여, 첫 번째 1-프래임이 카-프래임으로서 선택된다. 보다도 시퀀스 $\{t_0, t_1, \dots, t_n\}$ 가 주어지고, 여기서, $t_i \in V$ 1 번째 1-프래임일 때, $k_0, k_1 \in V$ 들은 카-프래임이라 하고, 레인지(range)가 그 이상에서 두 개의 영상들이 유사한 것으로 간주되지 않는 최대 거리를 한다.

$$\forall k_i \in V, s_i, \text{distance}(k_i, k_j) \leq d, \text{일때,}$$

$$S = \{s_i | s_i \leq s_i + N, N \geq 0\}$$

$$\exists s_i \in S, s_i, d = \text{distance}(k_i, k_j) \wedge \forall s_j \in S, s_j, d = \text{distance}(k_i, k_j)$$

마지막으로, 카-프래임 k_0 에 1-프래임 t_i 에 연관된 기술자 및 거리인 d_i 가 할당된다.

양호한 구현에서, N 개의 상이한 값수값들이 선택된다. 또한, 각 섷의 모든 1-프래임들이 이용된다(이 경우, N 은 섷들의 길이들에 의존함).

영상 유사성 검색이 가능하며, 단일 프로그램의 카-프래임들에 걸쳐 수행될 때 매우 효과적인 것으로 증명되었다. 상이한 프로그램들에서의 검색을 수행하는 것도 가능하지만, 높은 수의 카-프래임들이 포함되기 때문에 보다 느린 것이다. 수십인의 영상들에서, 두만 문자열들은 오만정(false positive)과 높은 시간이다.

오만정들은 유사한 시각적 기술자들을 가진 상이한 영상들로 인한 것이다. 영상들이 찾아낼 수 즉 유사한 영상들을 가지지만 예상한 상이한 프래임들을 발견할 가능성이 보다 높아진다. 오만정의 수를 줄이시거나, 각, 음역의 정밀도 증가시키기 위해서, 매우 효율적인 시각적 기술자들이 사용될 수 있다.

매우 신뢰가능한 기술자들은 그 연산적 복잡성으로 인해 응답 시간을 더 증가시킬 수 있다. 응답 시간을 감소시키기 위해서, 부가적 반복들이 취해질 수 있으며, 반복들이 사용될 수도 있다. 예를 들어, 레퍼터링을 사용하여 있는 것 만큼 적은 샘플링을 사용하여 유사 영상들의 첫 번째 세트를 선택하기 위해 거점(center) 기술자를 사용한다. 그후, 매우 광범위한고 연산적으로 고가의 기술자들만이 사용되어 세 번째 세트들의 영상들을 선택하고, 따라서, 보다 수렴가한 응답 시간을 필요로하게 된다. 두 번째 번째 기술자들의 빠르기로 순차적 소개를 취하는 것으로 이루어진다. R -트리, S -트리 또는 버트럼 같은 잘 기술 분야에 널리 공지되어 있는 데이터 액세스 구조들은 근세 데이터베이스를 형성하지 않고 단지 관련 영상들만을 보여줄 것을 가능하게 하는 방식으로 기술자들을 조직하는 것을 허용한다. 이들 데이터들은 기술자들이 보다 또는 빠른의 공간블록의 저장들로서 모으는 것을 필요로하며, 이들은 데이터베이스를 구성 시스템에 약간의 연산적 비용을 추가한다. 따라서, 이들은 매우 큰 데이터(영상) 데이터베이스들을 위해 적합하다.

본 발명을 그 양호한 실시예들을 참조로 설명하였지만, 이들은 비제한적인 예들이라는 것을 이해하여야 한

다. 따라서, 본 기술 분야의 숙련자들은 청구범위에 정의된 바와 같은 본 발명의 범주로부터 벗어나지 않고 다양한 변형들을 연출할 수 있다. 예로서, 기록물의 특징들을 포함하는 내용의 데이터들은 사용자에게 보다 매력적인 기록물을 제공하는 시계로서 제 2 위치에 의해 제공될 수 있다. 내용의 데이터들이 사용자의 관심사에 기반하는 경우, 그의 관심한 정보 신호의 부호를 사이의 유사성의 범위를 결정한다.

'포함하다'라는 용어와 그 속용형들은 청구범위에 기술된 것 이외의 다른 구성 요소들 또는 단계들의 존재를 배제하는 것은 아니다. 또한, 구성 요소 앞에 부속관사를 사용하는 것은 이런 구성 요소들에 복수개 존재하는 것을 배제하는 것은 아니다. 청구범위에서, 괄호를 사이에 넣거나 어떤 정보 부호들도 청구범위의 범주를 제한하는 것으로서 이해되어서는 안된다. 본 발명은 하드웨어 및 소프트웨어에 의해 구현될 수 있다. 다수의 '수단'은 하드웨어의 동등 아이템을 나타낼 수 있다. 부가적으로, 본 발명은 각각의 신규한 특징 또는 특징들의 조합으로 이루어진다.

(57) 청구의 범위

청구항 1

제 1 저장 매체상에 저장된 정보 신호를 재생하는 장치로서,

- 상기 제 1 저장 매체로부터 상기 정보 신호를 판독하는 판독 수단과,

- 상기 정보 신호를 디스플레이 유닛에 공급하는 출력 수단과,

- 사용자가 상기 정보 신호를 액세스할 수 있도록, 명령들을 수신하는 사용자 제어가능 입력 수단을 포함하는, 상기 정보 신호 재생 장치에 있어서,

상기 사용자 제어가능 입력 수단은 일련의 순간에 제 1 명령을 수신하도록 적응되고,

상기 장치는 상기 정보 신호내의 제 2 위치에서 상기 저장 매체로부터 상기 정보 신호를 판독하는 것을 시작하도록 상기 판독 수단을 제어하는 수단을 더 포함하며, 상기 제 2 위치에 있는 상기 정보 신호는 상기 제 1 명령을 수신하는 상기 순간에 판독된 제 1 위치에 있는 상기 정보 신호의 특징들, 또는, 상기 순간 이전에 판독된 정보 신호의 특징들과의 유사성을 나타내는 특징들을 가지는 것을 특징으로 하는, 정보 신호 재생 장치.

청구항 2

제 1 항에 있어서,

상기 제어 수단은 제 2 저장 매체로부터 데이터 신호를 판독하도록 더 적응되며,

상기 데이터 신호는 상기 정보 신호내의 위치를 및 유사한 특징들을 가진 정보 신호내의 다른 위치들에 대한 정보를 포함하는 것을 특징으로 하는, 정보 신호 재생 장치.

청구항 3

제 2 항에 있어서,

상기 정보 신호로부터 상기 특징들을 추출하는 추출 수단과,

상기 데이터 신호를 알도록 유사한 특징들을 가지는 위치들을 상기 추출된 특징들에 의존하여 결정하는 수단과,

상기 데이터 신호를 상기 제 2 저장 매체상에 기록하는 기록 수단을 더 포함하는 것을 특징으로 하는 정보 신호 재생 장치.

청구항 4

제 1 항에 있어서,

상기 특징들은 상기 정보 신호내의 위치들에 대응하는 영상들의 컬러 히스토그램과 관계를 가지는 것을 특징으로 하는, 정보 신호 재생 장치.

청구항 5

제 1 항에 있어서,

상기 특징들은 상기 정보 신호내의 위치들에 대응하는 영상들의 컬러 구조 히스토그램과 관계를 가지는 것을 특징으로 하는, 정보 신호 재생 장치.

청구항 6

제 1 항에 있어서,

상기 특징들은 상기 정보 신호내의 위치들에 대응하는 영상들의 컬러 구조 히스토그램과 관계를 가지는 것을 특징으로 하는, 정보 신호 재생 장치.

청구항 7

제 1 항에 있어서,

상기 사용자 제어가능 입력 수단은 상기 순간에 상기 정보 신호내의 위치에 후속하는 위치에 있는 상기 정보 신호를 판독하는 것을 시작하도록 상기 판독 수단을 제어하는 입력 수단을 포함하는 것을 특징으로 하는

는, 정보 신호 재생 장치.

종구형 8

제 1 항에 있어서,

상기 서술지 제2항에 기재된 입력 수단인 상기 순간에 상기 정보 신호내의 위치에 선행하는 위치에 있는 상기 정보 신호를 판독하는 것을 시작하도록 상기 판독 수단을 제어하는 입력 수단을 포함하는 것을 특징으로 하는, 정보 신호 재생 장치.

종구형 9

제 1 항에 기재된 장치에 저장된 정보 신호를 재생하는 방법으로서,

- 상기 제 1 저장 매체로부터 상기 정보 신호를 판독하는 단계와,

- 상기 정보 신호를 디스플레이 유닛에 공급하는 단계와,

- 사용자기 상기 정보 신호를 역재생할 수 있도록 명령들을 수신하는 단계를 포함하는, 상기 정보 신호 재생 방법에 있어서,

- 입력의 순서에 제 1 명령을 수신하는 단계와,

- 상기 정보 신호내의 제 2 위치에서 상기 저장매체로부터 상기 정보 신호를 판독하는 것을 시작하는 단계 (start reading)로서, 상기 제 2 위치에 있는 상기 정보 신호는 상기 제 1 명령을 수신하는 상기 순간에 판독된 제 1 위치에 있는 상기 정보 신호의 위치임을 나타내거나 또는, 상기 순간 이전에 판독된 상기 정보 신호의 부분과 유사성을 나타내는, 상기 정보 신호 판독 시작 단계를 더 포함하는 것을 특징으로 하는 정보 신호 재생 방법.

종구형 10

제 9 항에 따른 방법을 프로세서가 실행할 수 있도록 하는 컴퓨터 프로그램.

종구형 11

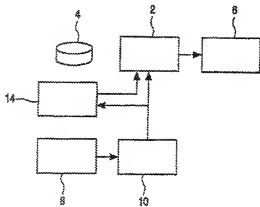
제 10 항에 따른 컴퓨터 프로그램을 담고 있는 유형 매체(tangible medium).

종구형 12

제 11 항에 따른 컴퓨터 프로그램을 담고 있는 신호.

도면

도면 1



도 B6P

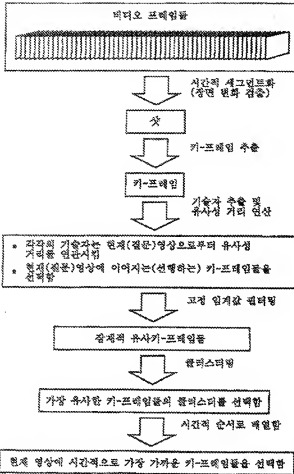


図 23

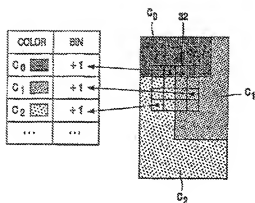


図 24

